

# A NOTE ON MARTINGALE HARDY SPACES

JINGGUO LAI

ABSTRACT. We investigate the relation between Carleson sequence and balayage, and use this to give an easy proof of the equivalence of the  $L^1$ -norms of the maximal function and the square function in non-homogeneous martingale settings.

## 1. INTRODUCTION AND THE MAIN THEOREM

In this note, we attempt to give an easy proof of the celebrated theorem:

*The  $L^1$ -norms of the maximal function and square function are equivalent.*

Throughout the note, we will work on the real line  $\mathbb{R}$  under Lebesgue measure, and assume our  $\sigma$ -algebras are generated by disjoint intervals. The notation  $\lesssim, \gtrsim$  stand for one-sided estimates up to an absolute constant, and the notation  $\approx$  stands for two-sided estimates up to an absolute constant. We will introduce the basic set-up following from [7].

*Definition 1.1.* A lattice  $\mathcal{L}$  is a collection of non-trivial finite intervals of  $\mathbb{R}$  with the following properties

- (i)  $\mathcal{L}$  is a union of *generations*  $\mathcal{L}_k, k \in \mathbb{Z}$ , where each generation is a collection of disjoint intervals, covering  $\mathbb{R}$ .
- (ii) For each  $k \in \mathbb{Z}$ , the covering  $\mathcal{L}_{k+1}$  is a finite refinement of the covering  $\mathcal{L}_k$ , i.e. each interval  $I \in \mathcal{L}_k$  is a finite union of disjoint intervals  $J \in \mathcal{L}_{k+1}$ . We allow the situation where there is only one such interval  $J$ , i.e.  $J = I$ ; this means that  $I \in \mathcal{L}_k$  also belongs to the generation  $\mathcal{L}_{k+1}$ .

We say a lattice  $\mathcal{L}$  is *homogenous* if

- (i) Each interval  $I \in \mathcal{L}_k$  is a union of *at most*  $r$  intervals  $J \in \mathcal{L}_{k+1}$ .
- (ii) There exists a constant  $K < \infty$  such that  $\frac{|I|}{|J|} \leq K$  for every  $I \in \mathcal{L}_k$  and every  $J \in \mathcal{L}_{k+1}, J \subseteq I$ .

The standard dyadic lattice  $\mathcal{D} = \{([0, 1) + j)2^k : j, k \in \mathbb{Z}\}$  is an example of homogenous lattice.

*Remark 1.2.* Let  $\mathfrak{A}_k$  be the  $\sigma$ -algebra generated by  $\mathcal{L}_k$ , i.e. countable unions of intervals in  $\mathcal{L}_k$ . Let  $\mathfrak{A}_{-\infty}$  be the largest  $\sigma$ -algebra contained in all  $\mathfrak{A}_k, k \in \mathbb{Z}$ , i.e.  $\mathfrak{A}_{-\infty} = \bigcap_{k \in \mathbb{Z}} \mathfrak{A}_k$ , and let  $\mathfrak{A}_{\infty}$  be the smallest  $\sigma$ -algebra containing all  $\mathfrak{A}_k, k \in \mathbb{Z}$ . The structures of  $\sigma$ -algebras  $\mathfrak{A}_{-\infty}$  and  $\mathfrak{A}_{\infty}$  can be described as follows

$\mathfrak{A}_{-\infty}$  is the  $\sigma$ -algebra generated by all the intervals  $I$  of form

$$I = \bigcup_{k \in \mathbb{Z}} I_k, \text{ where } I_k \in \mathcal{L}_k, I_k \subseteq I_{k-1}.$$

Note that  $\mathbb{R}$  is a disjoint union of such intervals  $I$  and at most countably many points (we might need to add left endpoints to the intervals  $I$ ). Let us denote the collection of such intervals  $I$  by  $\mathfrak{A}_{-\infty}^0$ . Define

$$\mathfrak{A}_{-\infty}^{0, \text{fin}} = \{I \in \mathfrak{A}_{-\infty}^0 : |I| < \infty\};$$

---

*Key words and phrases.* Carleson sequence, balayage, maximal function, square function.

'fin' here is to remind that the set consists of intervals of finite measure.

Instead of describing  $\mathfrak{A}_\infty$ , let us describe the corresponding measurable functions. *Namely, a function  $f(x)$  is  $\mathfrak{A}_\infty$ -measurable, if it is Borel measurable and it is constant on intervals  $I$ .*

$$I = \bigcap_{k \in \mathbb{Z}} I_k, \text{ where } I_k \in \mathcal{L}_k, I_k \subseteq I_{k-1}.$$

Clearly, such intervals  $I$  do not intersect, so there can only be countably many of them. In this note, we always assume that all functions are  $\mathfrak{A}_\infty$ -measurable.

*Definition 1.3.* The averaging operator  $\mathbb{E}_I$  is defined to be

$$\mathbb{E}_I f = \langle f \rangle_I \mathbf{1}_I = \left( \frac{1}{|I|} \int_I f \right) \mathbf{1}_I,$$

where  $\mathbf{1}_I$  is the indicator function of the interval  $I$ .

Let  $\mathbb{E}_k$  denote the conditional expectation

$$\mathbb{E}_k f = \mathbb{E}[f | \mathfrak{A}_k] = \sum_{I \in \mathcal{L}_k} \mathbb{E}_I f.$$

For an interval  $I \in \mathcal{L}$ , let  $\text{rk}(I)$  be the rank of the interval  $I$ , i.e. the largest number  $k$  such that  $I \in \mathcal{L}_k$ .

For an interval  $I \in \mathcal{L}$ ,  $\text{rk}(I) = k$ , a child of  $I$  is an interval  $J \in \mathcal{L}_{k+1}$  such that  $J \subseteq I$  (actually we can write  $J \subsetneq I$ ). The collection of all children of  $I$  is denoted by  $\text{child}(I)$ . Correspondingly,  $I$  is called the parent of  $J$ .

The difference operator  $\Delta_I$  is defined to be

$$\Delta_I f = -\mathbb{E}_I f + \sum_{J \in \text{child}(I)} \mathbb{E}_J f.$$

Let  $\Delta_k$  denote the martingale difference

$$\Delta_k f = \mathbb{E}_k f - \mathbb{E}_{k-1} f = \sum_{I \in \mathcal{L}, \text{rk}(I)=k-1} \Delta_I f.$$

To motivate the definition of the square function as well as for our later purpose, let us prove the following proposition:

**Proposition 1.4.** For a function  $f(x) \in L^p$ , where  $p > 1$ ,

$$(1.1) \quad f = \sum_{I \in \mathcal{L}} \Delta_I f + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \mathbb{E}_I f \quad \text{a.e. and in } L^p.$$

*Proof.* One can easily see that

$$\sum_{I \in \mathcal{L}, m \leq \text{rk}(I) < n} \Delta_I f = \sum_{m < k \leq n} \Delta_k f = \mathbb{E}_n f - \mathbb{E}_m f.$$

- $\mathbb{E}_n f \rightarrow f$  a.e. as  $n \rightarrow \infty$ .

For each  $x \in \mathbb{R}$ , consider the unique interval  $I_k(x)$  in  $\mathcal{L}_k$  containing  $x$ , and let  $I(x) = \bigcap_{k \in \mathbb{Z}} I_k(x) \in \mathfrak{A}_\infty$ .

If  $|I(x)| = 0$ , then by Lebesgue Differentiation Theorem,

$$\lim_{n \rightarrow \infty} \mathbb{E}_n f(x) = \lim_{n \rightarrow \infty} \frac{1}{|I_n(x)|} \int_{I_n(x)} f = f(x)$$

holds almost everywhere for such  $x$ .

If  $|I(x)| > 0$ , then since  $f$  is  $\mathfrak{A}_\infty$ -measurable, we have pointwisely,

$$\lim_{n \rightarrow \infty} \mathbb{E}_n f(x) = \lim_{n \rightarrow \infty} \frac{1}{|I_n(x)|} \int_{I_n(x)} f = \frac{1}{|I(x)|} \int_{I(x)} f = f(x).$$

- $\mathbb{E}_m f \rightarrow \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \mathbb{E}_I f$  as  $m \rightarrow -\infty$ .

For each  $x \in \mathbb{R}$ , take now  $I(x) = \bigcup_{k \in \mathbb{Z}} I_k(x)$ .

If  $|I(x)| = \infty$ , then since  $f \in L^p$ , we can estimate

$$|\mathbb{E}_m f(x)| \leq \left[ \frac{1}{|I_m(x)|} \int_{I_m(x)} |f|^p \right]^{\frac{1}{p}} \leq |I_m(x)|^{-\frac{1}{p}} \|f\|_p \rightarrow 0$$

as  $m \rightarrow -\infty$ .

If  $I(x) \in \mathfrak{A}_{-\infty}^{0, \text{fin}}$ , then pointwisely,

$$\lim_{m \rightarrow -\infty} \mathbb{E}_m f(x) = \lim_{m \rightarrow -\infty} \frac{1}{|I_m(x)|} \int_{I_m(x)} f = \frac{1}{|I(x)|} \int_{I(x)} f = \mathbb{E}_{I(x)} f(x).$$

- $f = \sum_{I \in \mathcal{L}} \Delta_I f + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \mathbb{E}_I f$  in  $L^p$ .

We have proved the almost everywhere convergence. Note that

$$|\mathbb{E}_n f - f| \leq 2f^* \quad \text{and} \quad \left| \mathbb{E}_m f - \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \mathbb{E}_I f \right| \leq 2f^*,$$

where  $f^*$  is the Hardy-Littlewood maximal function of  $f$ . Hardy-Littlewood Maximal Inequality and Dominated Convergence Theorem imply the  $L^p$  convergence.  $\square$

*Definition 1.5.* The *maximal function* is defined to be

$$Mf(x) = \sup_{x \in I, I \in \mathcal{L}} |\mathbb{E}_I f| = \sup_{x \in I, I \in \mathcal{L}} \frac{1}{|I|} \left| \int_I f \right|.$$

We say that  $f \in H_*^1$  if  $Mf \in L^1$  with norm  $\|f\|_{H_*^1} = \|Mf\|_1$ .

The *square function* is defined to be

$$Sf(x) = \left[ \sum_{I \in \mathcal{L}} |\Delta_I f|^2 + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} |\mathbb{E}_I f|^2 \right]^{\frac{1}{2}}.$$

We say that  $f \in H_S^1$  if  $Sf \in L^1$  with norm  $\|f\|_{H_S^1} = \|Sf\|_1$ .

Both  $H_*^1$  and  $H_S^1$  are called *martingale Hardy spaces*.

Now we are ready to state our main theorem:

**Theorem 1.6** (B. J. Davis [1]).  $\|Mf\|_1 \approx \|Sf\|_1$  or  $H_*^1 = H_S^1$ .

This theorem and its proof can be found for example in [1], [2], and [3], all using probabilistic method. We will give a relatively easier analysis proof here, in the spirit of [3], but with the help of Carleson sequence and balayage, which is also interesting by itself.

## 2. CARLESON SEQUENCE AND BALAYAGE

This section is devoted to investigating the relation between Carleson sequence and balayage. We start with the following definitions:

*Definition 2.1.* A sequence of non-negative numbers  $\{a_I\}_{I \in \mathcal{L}}$  is a *Carleson sequence*, if there exists a constant  $C > 0$  such that

$$\frac{1}{|I|} \sum_{J \in \mathcal{L}, J \subseteq I} a_J \leq C$$

for each  $I \in \mathcal{L} \cup \mathfrak{A}_{-\infty}^{0, \text{fin}}$ . Moreover, we denote the infimum of such constant by  $\text{Carl}(a_I)$ .

*Definition 2.2.* Given a sequence of numbers  $\{a_I\}_{I \in \mathcal{L}}$ , its *balayage* is defined to be the function

$$g(x) = \sum_{I \in \mathcal{L}} \frac{1}{|I|} \mathbf{1}_I a_I.$$

*Definition 2.3.* A locally integrable function  $g(x)$  on  $\mathbb{R}$  is in the class of *martingale BMO*, if

- $\left[ \frac{1}{|I|} \int_I \sum_{J \in \mathcal{L}, J \subseteq I} |\Delta_J g|^2 \right]^{\frac{1}{2}} = C_1 < \infty$ , for all  $I \in \mathcal{L}$ .
- $\|\Delta_I g\|_{\infty} = C_2 < \infty$ , for all  $I \in \mathcal{L}$ .

Moreover, we define  $\|g\|_{\text{BMO}} = \max\{C_1, C_2\}$ .

*Remark 2.4.* As in the classical case, martingale BMO space does not distinguish constant functions. In other words, we should think of martingale BMO space as the quotient of the above class by the space of  $\{g \in L_{loc}^1 : \mathbb{E}_I g \text{ is constant, for all } I \in \mathfrak{A}_{-\infty}^0\}$ .

*Remark 2.5.* This definition of martingale BMO functions is stronger than the classical one in harmonic analysis, since by (1.1) and orthogonality of the difference operators,

$$(2.1) \quad \frac{1}{|I|} \int_I |g(x) - \langle g \rangle_I|^2 = \frac{1}{|I|} \int_I \lim_{n \rightarrow \infty} \left| \sum_{J \in \mathcal{L}, J \subseteq I, \text{rk}(I) < n} \Delta_J g \right|^2 = \frac{1}{|I|} \int_I \sum_{J \in \mathcal{L}, J \subseteq I} |\Delta_J g|^2,$$

where  $\langle g \rangle_I = \frac{1}{|I|} \int_I g$  is the average of  $g(x)$  over  $I$ .

It was known for a long time that the additional condition is needed. One can look, for example, at [2] in detail.

Now we have the precise statement:

**Theorem 2.6.** (Carleson sequence and balayage)

- (i) Given a Carleson sequence  $\{|a_I|\}_{I \in \mathcal{L}}$ , the following inequality holds

$$(2.2) \quad \left\| g(x) = \sum_{I \in \mathcal{L}} \frac{1}{|I|} \mathbf{1}_I a_I \right\|_{\text{BMO}} \leq 2 \text{Carl}(|a_I|).$$

- (ii) For each function  $g(x)$  in the class of martingale BMO, there exists  $\phi(x) \in L^\infty$ , and a Carleson sequence  $\{|a_I|\}_{I \in \mathcal{L}}$ , such that

$$(2.3) \quad g = \phi + \sum_{I \in \mathcal{L}} \frac{1}{|I|} \mathbf{1}_I a_I,$$

where  $\|\phi\|_{\infty} \leq 2\|g\|_{\text{BMO}}$ , and  $\text{Carl}(|a_I|) \leq 3\|g\|_{\text{BMO}}$ .

*Remark 2.7.* One might hope that the reverse inequality to (2.2) holds in the strict sense, but this is not the case due to [6]. Nevertheless, we still have (2.3) as a partial reverse answer, which is good enough for our purpose.

*Proof.* (i) Check by definition

$$\begin{aligned}
\frac{1}{|I|} \int_I \sum_{J \in \mathcal{L}, J \subseteq I} |\Delta_J g|^2 &= \frac{1}{|I|} \int_I |g(x) - \langle g \rangle_I|^2 \quad (2.1) \\
&= \langle g^2 \rangle_I - \langle g \rangle_I^2 \\
&= \frac{1}{|I|} \int_I \left| \sum_{J \in \mathcal{L}} \frac{1}{|J|} \mathbf{1}_J a_J \right|^2 - \left( \frac{1}{|I|} \int_I \sum_{J \in \mathcal{L}} \frac{1}{|J|} \mathbf{1}_J a_J \right)^2 \\
&= \frac{1}{|I|} \int_I \left| \sum_{J \in \mathcal{L}, J \subseteq I} \frac{1}{|J|} \mathbf{1}_J a_J \right|^2 - \left( \frac{1}{|I|} \int_I \sum_{J \in \mathcal{L}, J \subseteq I} \frac{1}{|J|} \mathbf{1}_J a_J \right)^2 \\
&\leq \frac{1}{|I|} \int_I \left( \sum_{J \in \mathcal{L}, J \subseteq I} \frac{1}{|J|} \mathbf{1}_J |a_J| \right)^2 \\
&= \frac{1}{|I|} \int_I \sum_{J, K \subseteq I} \frac{1}{|J|} \frac{1}{|K|} |a_J| |a_K| \mathbf{1}_J \mathbf{1}_K \\
&= \frac{1}{|I|} \sum_{J \subseteq I} |a_J| \frac{1}{|J|} \sum_{K \subseteq J} |a_K| + \frac{1}{|I|} \sum_{K \subseteq I} |a_K| \frac{1}{|K|} \sum_{J \not\subseteq K} |a_J| \\
&\leq \frac{1}{|I|} \sum_{J \subseteq I} |a_J| \text{Carl}(|a_I|) + \frac{1}{|I|} \sum_{K \subseteq I} |a_K| \text{Carl}(|a_I|) \\
&\leq 2 \text{Carl}(|a_I|)^2,
\end{aligned}$$

holds for any  $I \in \mathcal{L}$ , where the forth equality follows from

$$\sum_{J \in \mathcal{L}, J \not\supseteq I} \frac{1}{|J|} \mathbf{1}_J a_J$$

being a constant on  $I$ .

Moreover, we can compute that for any  $I \in \mathcal{L}$ ,

$$\begin{aligned}
\mathbb{E}_I g &= \langle g \rangle_I \mathbf{1}_I = \left( \frac{1}{|I|} \int_I \sum_{K \in \mathcal{L}} \frac{1}{|K|} \mathbf{1}_K a_K \right) \mathbf{1}_I \\
&= \left( \frac{1}{|I|} \sum_{K \in \mathcal{L}, K \subseteq I} a_K \right) \mathbf{1}_I + \left( \sum_{K \in \mathcal{L}, K \not\subseteq I} \frac{1}{|K|} a_K \right) \mathbf{1}_I.
\end{aligned}$$

Thus, we can estimate

$$\begin{aligned}
|\Delta_I g| &= | -\mathbb{E}_I g + \sum_{J \in \text{child}(I)} \mathbb{E}_J g | \\
&= \left| \sum_{J \in \text{child}(I)} \left( \frac{1}{|J|} \sum_{K \in \mathcal{L}, K \subseteq J} a_K \right) \mathbf{1}_J + \sum_{J \in \text{child}(I)} \left( \sum_{K \in \mathcal{L}, K \not\subseteq J} \frac{1}{|K|} a_K \right) \mathbf{1}_J \right. \\
&\quad \left. - \left( \frac{1}{|I|} \sum_{K \in \mathcal{L}, K \subseteq I} a_K \right) \mathbf{1}_I - \left( \sum_{K \in \mathcal{L}, K \not\subseteq I} \frac{1}{|K|} a_K \right) \mathbf{1}_I \right|
\end{aligned}$$

$$\begin{aligned}
&= \left| \sum_{J \in \text{child}(I)} \left( \frac{1}{|J|} \sum_{K \in \mathcal{L}, K \subseteq J} a_K \right) \mathbf{1}_J - \left( \frac{1}{|I|} \sum_{K \in \mathcal{L}, K \subseteq I} a_K \right) \mathbf{1}_I \right| \\
&\leq \sum_{J \in \text{child}(I)} \left( \frac{1}{|J|} \sum_{K \in \mathcal{L}, K \subseteq J} |a_K| \right) \mathbf{1}_J + \left( \frac{1}{|I|} \sum_{K \in \mathcal{L}, K \subseteq I} |a_K| \right) \mathbf{1}_I \leq 2\text{Carl}(|a_I|).
\end{aligned}$$

- (ii) The proof of this statement is a modification of the proof of John-Nirenberg theorem which is suggested in [5].

Given a martingale BMO function  $g(x)$ , by (2.1) we have

$$\sup_{I \in \mathcal{L}} \frac{1}{|I|} \int_I |g(x) - \langle g \rangle_I| \leq \sup_{I \in \mathcal{L}} \left[ \frac{1}{|I|} \int_I |g(x) - \langle g \rangle_I|^2 \right]^{\frac{1}{2}} = \|g\|_{\text{BMO}}.$$

Fix the 0-th generation  $\mathcal{L}_0$  of the lattice  $\mathcal{L}$ , recall that  $\mathcal{L}_0$  is a collection of non-trivial, finite, disjoint intervals covering  $\mathbb{R}$ , so we can write  $\mathcal{L}_0 = \{I_m\}_{m \in \mathfrak{M}}$ .

Over each interval  $I_m, m \in \mathfrak{M}$ , let us first do the Calderón-Zygmund decomposition to the function  $|g(x) - \langle g \rangle_{I_m}|$  at height  $\lambda = 2\|g\|_{\text{BMO}}$  with respect to the lattice  $\mathcal{L}$ , from this we obtain:

A collection of disjoint intervals  $\{I_j^{(1)}\} \subseteq \mathcal{L}$  contained in  $I_m$ , such that over each  $I_j^{(1)}$ ,

$$\frac{1}{|I_j^{(1)}|} \int_{I_j^{(1)}} |g(x) - \langle g \rangle_{I_m}| > 2\|g\|_{\text{BMO}}.$$

This immediately implies

- $|g(x) - \langle g \rangle_{I_m}| \leq 2\|g\|_{\text{BMO}}$  a.e. over  $I_m \setminus \bigcup_j I_j^{(1)}$ .

For each  $x \in I_m \setminus \bigcup_j I_j^{(1)}$ , from the Calderón-Zygmund decomposition, we know that over all intervals  $\{I_k(x)\}_{k \geq 1}$ ,

$$\frac{1}{|I_k(x)|} \int_{I_k(x)} |g(x) - \langle g \rangle_{I_m}| \leq 2\|g\|_{\text{BMO}},$$

where  $I_k(x)$  is the unique interval in  $\mathcal{L}_k, k \geq 1$  containing  $x$ .

If  $I(x) = \bigcap_{k \geq 1} I_k(x) \in \mathfrak{A}_\infty$  has  $|I(x)| = 0$ , then by Lebesgue Differentiation Theorem, we can conclude that

$$|g(x) - \langle g \rangle_{I_m}| = \lim_{k \rightarrow \infty} \frac{1}{|I_k(x)|} \int_{I_k(x)} |g(x) - \langle g \rangle_{I_m}| \leq 2\|g\|_{\text{BMO}}$$

holds almost everywhere for such  $x$ .

If  $|I(x)| > 0$ , then since  $g(x)$  is  $\mathfrak{A}_\infty$ -measurable, we can conclude that

$$\begin{aligned}
|g(x) - \langle g \rangle_{I_m}| &= \frac{1}{|I(x)|} \int_{I(x)} |g(x) - \langle g \rangle_{I_m}| \\
&= \lim_{k \rightarrow \infty} \frac{1}{|I_k(x)|} \int_{I_k(x)} |g(x) - \langle g \rangle_{I_m}| \leq 2\|g\|_{\text{BMO}}.
\end{aligned}$$

- $\sum_j |I_j^{(1)}| \leq \frac{1}{2}|I_m|$ .

$$\begin{aligned}
\sum_j |I_j^{(1)}| &\leq \sum_j \frac{1}{2\|g\|_{\text{BMO}}} \int_{I_j^{(1)}} |g(x) - \langle g \rangle_{I_m}| \\
&\leq \frac{1}{2\|g\|_{\text{BMO}}} \int_{I_m} |g(x) - \langle g \rangle_{I_m}| \leq \frac{1}{2}|I_m|.
\end{aligned}$$

Next, let us do the Calderón-Zygmund decomposition over each  $I_j^{(1)}$  to the function  $|g(x) - \langle g \rangle_{I_j^{(1)}}|$  at again height  $\lambda = 2\|g\|_{\text{BMO}}$  with respect to the lattice  $\mathcal{L}$ , from this we obtain

A collection of disjoint intervals  $\{I_k^{(2)}\} \subseteq \mathcal{L}$  contained in  $I_j^{(1)}$ , such that over each  $I_k^{(2)}$ ,

$$\frac{1}{|I_k^{(2)}|} \int_{I_k^{(2)}} |g(x) - \langle g \rangle_{I_j^{(1)}}| > 2\|g\|_{\text{BMO}}.$$

Similar argument yields

- $|g(x) - \langle g \rangle_{I_j^{(1)}}| \leq 2\|g\|_{\text{BMO}}$  a.e. over  $I_j^{(1)} \setminus \bigcup_k I_k^{(2)}$ .
- $\sum_k |I_k^{(2)}| \leq (\frac{1}{2})|I_j^{(1)}|$ .

Moreover, if we sum over all intervals  $\{I_j^{(1)}\}$ , we will have

$$\sum_k |I_k^{(2)}| \leq \frac{1}{2} \sum_j |I_j^{(1)}| \leq (\frac{1}{2})^2 |I_m|.$$

Continue this process indefinitely. At stage  $n$  we get disjoint intervals  $\{I_k^{(n)}\} \subseteq \mathcal{L}$ , such that over each  $I_k^{(n)}$ ,

- $|g(x) - \langle g \rangle_{I_j^{(n-1)}}| \leq 2\|g\|_{\text{BMO}}$  a.e. over  $I_j^{(n-1)} \setminus \bigcup_k I_k^{(n)}$ .
- $\sum_k |I_k^{(n)}| \leq (\frac{1}{2})^n |I_m|$ .

Now take  $\mathcal{G}_m = \bigcup_{j,n} \{I_j^{(n)}\} \cup \{I_m\}$ , and let  $\mathcal{G} = \bigcup_m \mathcal{G}_m \subseteq \mathcal{L}$ , define  $\{a_I\}_{I \in \mathcal{L}}$  to be

$$a_{I_k^{(n)}} = |I_k^{(n)}| (\langle g \rangle_{I_k^{(n)}} - \langle g \rangle_{I_j^{(n-1)}}),$$

and for  $I \notin \mathcal{G}$ , simply set

$$a_I = 0.$$

Finally, define the function  $\phi(x) = g(x) - \sum_{I \in \mathcal{L}} \frac{1}{|I|} \mathbf{1}_I a_I$ .

*Claim 1:*  $\|\phi\|_\infty \leq 2\|g\|_{\text{BMO}}$ .

If there exists  $x \in I_m, I_m \in \mathcal{L}_0$ , such that at all stages  $n \geq 1$ , one can find an interval  $I_k^{(n)}$  containing  $x$ , let  $I(x) \in \mathfrak{A}_\infty$  be the intersection of all those intervals, then by construction

$$|\phi(x)| = |g(x) - \sum_{I \in \mathcal{L}} \frac{1}{|I|} \mathbf{1}_I a_I| = |g(x) - \lim_{n \rightarrow \infty} \langle g \rangle_{I_k^{(n)}}| = |g(x) - \langle g \rangle_{I(x)}| = 0.$$

Otherwise, for arbitrary  $I_m \in \mathcal{L}_0$ , we have

$$|\phi(x)| = |g(x) - \sum_{I \in \mathcal{L}} \frac{1}{|I|} \mathbf{1}_I a_I| = |g(x) - \langle g \rangle_{I_j^{(n-1)}}| \leq 2\|g\|_{\text{BMO}}$$

almost everywhere over  $I_j^{(n-1)} \setminus \bigcup_k I_k^{(n)}$  for all stages  $n \geq 1$ , therefore  $\|\phi\|_\infty \leq 2\|g\|_{\text{BMO}}$ .

*Claim 2:*  $\text{Carl}(|a_I|) \leq 3\|g\|_{\text{BMO}}$ .

Fix an interval  $J \in \mathcal{L} \cup \mathfrak{A}_{-\infty}^{0, \text{fin}}$ , we have

$$\frac{1}{|J|} \sum_{I \in \mathcal{L}, I \subseteq J} |a_I| = \frac{1}{|J|} \sum_{I_k^{(n)} \subseteq J} |I_k^{(n)}| |\langle g \rangle_{I_k^{(n)}} - \langle g \rangle_{I_j^{(n-1)}}|.$$

Note that  $g \in BMO$  gives  $\|\Delta_I g\|_\infty \leq \|g\|_{BMO}$ , for all  $I \in \mathcal{L}$ , so

$$\begin{aligned} |\langle g \rangle_{I_k^{(n)}} - \langle g \rangle_{I_j^{(n-1)}}| &\leq |\langle g \rangle_{I_k^{(n)}} - \langle g \rangle_{\widetilde{I_k^{(n)}}}| + |\langle g \rangle_{\widetilde{I_k^{(n)}}} - \langle g \rangle_{I_j^{(n-1)}}| \\ &\leq \|\Delta_{\widetilde{I_k^{(n)}}} g\|_\infty + \frac{1}{|\widetilde{I_k^{(n)}}|} \int_{\widetilde{I_k^{(n)}}} |g(x) - \langle g \rangle_{I_j^{(n-1)}}| \\ &\leq \|g\|_{BMO} + 2\|g\|_{BMO} = 3\|g\|_{BMO}, \end{aligned}$$

where  $\widetilde{I_k^{(n)}}$  is the parent of  $I_k^{(n)}$ .

To estimate  $\frac{1}{|J|} \sum_{I_k^{(n)} \subseteq J} |I_k^{(n)}|$ , by our construction, it suffices to assume  $J$  being contained in some  $I_m \in \mathcal{L}_0$ , and

$$\frac{1}{|J|} \sum_{I_k^{(n)} \subseteq J} |I_k^{(n)}| = \frac{1}{|J|} \sum_{n \geq 1} \sum_k |I_k^{(n)}| \leq \frac{1}{|J|} \sum_{n \geq 1} \left(\frac{1}{2}\right)^n |J| = 1.$$

Hence, we obtain  $\frac{1}{|J|} \sum_{I \in \mathcal{L}, I \subseteq J} |a_I| \leq 3\|g\|_{BMO}$ , for all  $J \in \mathcal{L} \cup \mathfrak{A}_{-\infty}^{0, \text{fin}}$ , which completes the proof.  $\square$

### 3. PROOF OF THE MAIN THEOREM

We are now ready to prove **Theorem 1.6** using **Theorem 2.6**.

*Proof.* (i)  $\|Mf\|_1 \lesssim \|Sf\|_1$  or  $H_*^1 \supseteq H_S^1$

To prove this half of the theorem, we need the following celebrated Fefferman's inequality [4] whose proof will be postponed to the next section:

**Theorem 3.1.** *For  $f \in H_S^1$  and  $g \in BMO$ , we have*

$$(3.1) \quad \int fg \leq 2\|f\|_{H_S^1} \|g\|_{BMO}.$$

Now, define the 'cut-off' of  $Mf(x)$  by

$$M_n f(x) = \max_{x \in I, I \in \bigcup_{k=-n}^n \mathcal{L}_k} |\langle f \rangle_I|,$$

so obviously,  $M_n f$  increases to  $Mf$  pointwisely. Moreover, for any  $J \in \mathcal{L}_n$ , we have  $M_n f|_J$  must be constant. In addition, define

$$I(J) = \{I \supseteq J, I \text{ is maximal in } \bigcup_{k=-n}^n \mathcal{L}_k : M_n f|_J = |\langle f \rangle_I|\},$$

and let  $a_{I(J)}$  be such that  $a_{I(J)} \langle f \rangle_{I(J)} = \int_J M_n f$ , which implies  $|a_{I(J)}| = |J|$ . Finally define

$$a_I = \sum_{\{J: I(J)=I\}} a_{I(J)},$$



therefore, we have

$$\langle f \rangle_I a_I = \langle f \rangle_I \left[ \sum_{\{J: I(J)=I\}} a_{I(J)} \right] = \sum_{\{J: I(J)=I\}} \langle f \rangle_{I(J)} a_{I(J)} = \sum_{\{J: I(J)=I\}} \int_J M_n f.$$

Summing over all intervals  $I \in \bigcup_{k=-n}^n \mathcal{L}_k$ , we obtain

$$\sum_{I \in \bigcup_{k=-n}^n \mathcal{L}_k} \langle f \rangle_I a_I = \sum_{I \in \bigcup_{k=-n}^n \mathcal{L}_k} \sum_{\{J: I(J)=I\}} \int_J M_n f = \int M_n f.$$

Having this, we can write

$$\int M_n f = \sum_{I \in \bigcup_{k=-n}^n \mathcal{L}_k} \langle f \rangle_I a_I = \int f \sum_{I \in \bigcup_{k=-n}^n \mathcal{L}_k} \frac{1}{|I|} \mathbf{1}_I a_I = \int f g,$$

where  $g(x) = \sum_{I \in \bigcup_{k=-n}^n \mathcal{L}_k} \frac{1}{|I|} \mathbf{1}_I a_I$ .

*Claim:*  $\|g\|_{\text{BMO}} \leq 2$ .

Let us first consider  $\text{Carl}(|a_I|)$  for the sequence  $\{|a_I|\}$ , such that  $I \in \bigcup_{k=-n}^n \mathcal{L}_k$ . Take any  $K \in \bigcup_{k=-n}^n \mathcal{L}_k$ , we have

$$\sum_{I \subseteq K} |a_I| = \sum_{I \subseteq K} \left| \sum_{\{J: I(J)=I\}} a_{I(J)} \right| \leq \sum_{J \in \mathcal{L}_n, J \subseteq K} |a_{I(J)}| = \sum_{J \in \mathcal{L}_n, J \subseteq K} |J| \leq |K|,$$

so  $\text{Carl}(|a_I|) \leq 1$ , and by (2.2), we know  $\|g\|_{\text{BMO}} \leq 2$ .

To complete our proof, by (3.1), we can conclude

$$\int M_n f = \int f g \leq 2 \|f\|_{H_S^1} \|g\|_{\text{BMO}} \leq 4 \|f\|_{H_S^1}.$$

Letting  $n \rightarrow \infty$ , by Monotone Convergence Theorem, we obtain

$$\|Mf\|_1 \leq 4 \|Sf\|_1.$$

(ii)  $\|Mf\|_1 \gtrsim \|Sf\|_1$  or  $H_*^1 \subseteq H_S^1$

To prove this half of the theorem, we need another auxiliary lemma which will also be proved in the next section:

**Lemma 3.2.** *For  $f \in H_*^1$  and a Carleson sequence  $\{|a_I|\}_{I \in \mathcal{L}}$ , we have*

$$(3.2) \quad \left| \sum_{I \in \mathcal{L}} \langle f \rangle_I a_I \right| \leq \|f\|_{H_*^1} \text{Carl}(|a_I|).$$

By (2.3) and (3.2), we can conclude that fix a function  $f \in H_*^1$ , and for any  $g \in \text{BMO}$ , we have

$$\begin{aligned} \left| \int f g \right| &= \left| \int f \left( \phi + \sum_{I \in \mathcal{L}} \frac{1}{|I|} \mathbf{1}_I a_I \right) \right| \leq \left| \int f \phi \right| + \left| \int f \sum_{I \in \mathcal{L}} \frac{1}{|I|} \mathbf{1}_I a_I \right| \\ &\leq \left| \int Mf \phi \right| + \left| \sum_{I \in \mathcal{L}} \langle f \rangle_I a_I \right| \leq \|Mf\|_1 \|\phi\|_\infty + \|f\|_{H_*^1} \text{Carl}(|a_I|) \\ &\leq \|f\|_{H_*^1} 2 \|g\|_{\text{BMO}} + \|f\|_{H_*^1} 3 \|g\|_{\text{BMO}} = 5 \|f\|_{H_*^1} \|g\|_{\text{BMO}}. \end{aligned}$$

*Claim:* The above estimation implies  $\|Sf\|_1 \lesssim \|Mf\|_1$ .

It suffices to prove this claim on a dense set of functions  $f(x)$  for which the corresponding sequence  $\{\Delta_I f\}_{I \in \mathcal{L}}$  and  $\{\mathbb{E}_I f\}_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}}$  have only finitely many non-zero terms.

Let us construct a function  $g \in BMO$  from the function  $f$ , such that  $\int Sf = \int fg$ .

First take  $g_I = \Delta_I f / Sf$  for  $I \in \mathcal{L}$ , and take  $g_I = \mathbb{E}_I f / Sf$  for  $I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}$ , thus  $\sum_{I \in \mathcal{L}} |g_I|^2 + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} |g_I|^2 = 1$ , and  $\sum_{I \in \mathcal{L}} \Delta_I f g_I + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \mathbb{E}_I f g_I = Sf$ . One can also check easily that

$$\begin{aligned} \int Sf &= \sum_{I \in \mathcal{L}} \int \Delta_I f g_I + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \int \mathbb{E}_I f g_I \\ &= \sum_{I \in \mathcal{L}} \int \Delta_I f \Delta_I g_I + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \int \mathbb{E}_I f \mathbb{E}_I g_I. \end{aligned}$$

Consider now  $g = \sum_{I \in \mathcal{L}} \Delta_I g_I + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \mathbb{E}_I g_I$ , since  $\{\Delta_I f\}_{I \in \mathcal{L}}$  and  $\{\mathbb{E}_I f\}_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}}$  have only finitely many non-zero terms, this is actually a finite sum which is obviously well-defined. Moreover, it is easy to compute that  $\Delta_I g = \Delta_I g_I$  for all  $I \in \mathcal{L}$ , and  $\mathbb{E}_I g = \mathbb{E}_I g_I$  for all  $I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}$ . Let us show by definition that  $g \in BMO$ .

For any interval  $I \in \mathcal{L}$ ,  $|\Delta_I g| = |\Delta_I g_I| \leq 2g_I^*$  where  $g_I^*$  is the Hardy-Littlewood maximal function of  $g_I$ . Note that  $\|g_I\|_\infty \leq 1$  for all  $I \in \mathcal{L}$ , thus we always have  $g_I^* \leq 1$  and so  $|\Delta_I g| \leq 2$ .

For any interval  $I \in \mathcal{L}$ , by Hardy-Littlewood Maximal Inequality, we can estimate

$$\int_I \sum_{J \subseteq I} |\Delta_J g|^2 = \int_I \sum_{J \subseteq I} |\Delta_J g_J|^2 \leq 4 \int_I \sum_{J \subseteq I} |g_J^*|^2 \leq C \int_I \sum_{J \subseteq I} |g_J|^2 \leq C|I|.$$

Hence, we find a function  $g \in BMO$  such that

$$\begin{aligned} \int Sf &= \sum_{I \in \mathcal{L}} \int \Delta_I f \Delta_I g_I + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \int \mathbb{E}_I f \mathbb{E}_I g_I \\ &= \sum_{I \in \mathcal{L}} \int \Delta_I f \Delta_I g + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \int \mathbb{E}_I f \mathbb{E}_I g \\ &= \int \left( \sum_{I \in \mathcal{L}} \Delta_I f + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \mathbb{E}_I f \right) \left( \sum_{I \in \mathcal{L}} \Delta_I g + \sum_{I \in \mathfrak{A}_{-\infty}^{0, \text{fin}}} \mathbb{E}_I g \right) = \int fg. \end{aligned}$$

Finally, we can conclude

$$\|f\|_{H_S^1} = \int Sf = \int fg \leq 5\|f\|_{H_*^1} \|g\|_{BMO} \lesssim \|f\|_{H_*^1}.$$

□

#### 4. PROOF OF THE AUXILIARY LEMMAS

In this section, we will prove **Theorem 3.1** and **Lemma 3.2**, and hence complete the whole proof.

To show Fefferman's inequality, we take the elegant proof from [3].

*Proof. Theorem 3.1*

For  $f \in H_S^1$  and  $g \in BMO$ , we can assume by Remark 2.4 that  $\mathbb{E}_I g = 0$ , for all  $I \in \mathfrak{A}_\infty^0$ , and so we can compute

$$|\int f g| = |\int (\sum_{I \in \mathcal{L}} \Delta_I f) (\sum_{I \in \mathcal{L}} \Delta_I g)| = |\sum_{I \in \mathcal{L}} \int \Delta_I f \Delta_I g| = |\sum_{k \in \mathbb{Z}} \int \Delta_k f \Delta_k g|.$$

Now, let us define  $S_n f(x) = \left[ \sum_{k \leq n} |\Delta_k f|^2 \right]^{\frac{1}{2}} = \left[ \sum_{\text{rk}(I) \leq n-1} |\Delta_I f|^2 \right]^{\frac{1}{2}}$ , then  $0 \leq S_n f(x) \leq S_{n+1} f(x) \leq S f(x)$  holds for all  $x$ . A very clever idea due to C. Herz suggests that

$$|\sum_{k \leq n} \int \Delta_k f \Delta_k g| = |\sum_{k \leq n} \int \frac{\Delta_k f}{\sqrt{S_k f}} \Delta_k g \sqrt{S_k f}| \leq \left( \sum_{k \leq n} \int \frac{|\Delta_k f|^2}{S_k f} \right)^{\frac{1}{2}} \left( \sum_{k \leq n} \int |\Delta_k g|^2 S_k f \right)^{\frac{1}{2}}.$$

Therefore, all boil down to the following simple estimations

$$\sum_{k \leq n} \int \frac{|\Delta_k f|^2}{S_k f} = \sum_{k \leq n} \int \frac{S_k^2 f - S_{k-1}^2 f}{S_k f} \leq 2 \sum_{k \leq n} \int (S_k f - S_{k-1} f) \leq 2 \int S f,$$

and

$$\sum_{k \leq n} \int |\Delta_k g|^2 S_k f = \sum_{k \leq n} \int \left( \sum_{l=k}^n |\Delta_l g|^2 \right) (S_k f - S_{k-1} f),$$

note that for every  $k$ ,  $S_k f - S_{k-1} f$  is a constant on children of intervals  $I$  such that  $\text{rk}(I) = k-1$  and equals 0 outside intervals  $I$ , thus for an interval  $J \in \text{child}(I)$ , we have

$$\begin{aligned} \int_J \left( \sum_{l=k}^n |\Delta_l g|^2 \right) (S_k f - S_{k-1} f) &= \int_J \left[ \frac{1}{|J|} \int_J \sum_{l=k}^n |\Delta_l g|^2 \right] (S_k f - S_{k-1} f) \\ &\leq \int_J \left[ \frac{1}{|J|} \int_J |\Delta_I g|^2 + \sum_{K \in \mathcal{L}, K \subseteq J} |\Delta_K g|^2 \right] (S_k f - S_{k-1} f) \\ &\leq 2 \|g\|_{BMO}^2 \int_J (S_k f - S_{k-1} f). \quad (\text{by Definition 2.3}) \end{aligned}$$

This given, we can estimate

$$\sum_{k \leq n} \int \left( \sum_{l=k}^n |\Delta_l g|^2 \right) (S_k f - S_{k-1} f) \leq 2 \|g\|_{BMO}^2 \sum_{k \leq n} \int (S_k f - S_{k-1} f) \leq 2 \|g\|_{BMO}^2 \int S f.$$

□

To show the second auxiliary lemma, we modify the proof using level sets comparision in [7] to get this sharper inequality.

*Proof. Lemma 3.2*

Define  $E_k = \{x \in \mathbb{R} : Mf(x) > (1+\varepsilon)^k\}$ , and take  $\mathcal{E}_k = \{I \in \mathcal{L} : I \in E_k\}$ . Note that  $E_k$  is a finite disjoint union of maximal intervals in  $\mathcal{E}_k$ , maximal is considered in the sense of inclusion.

*Claim 1:*  $|\sum_{I \in \mathcal{L}} \langle f \rangle_I a_I| \leq \varepsilon \sum_{k \in \mathbb{Z}} (1+\varepsilon)^k |E_k| \text{Carl}(|a_I|).$

First note that  $\sum_{I \in \mathcal{E}_k} |a_I| \leq \sum_{I \subseteq E_k} |a_I| \leq |E_k| \text{Carl}(|a_I|)$ , because  $\{|a_I|\}_{I \in \mathcal{L}}$  is a Carleson sequence. Using this fact, we can estimate

$$\begin{aligned}
\varepsilon \sum_{k \in \mathbb{Z}} (1 + \varepsilon)^k |E_k| \text{Carl}(|a_I|) &\geq \varepsilon \sum_{k \in \mathbb{Z}} (1 + \varepsilon)^k \sum_{I \in \mathcal{E}_k} |a_I| = \varepsilon \sum_{k \in \mathbb{Z}} (1 + \varepsilon)^k \sum_{l=0}^{\infty} \sum_{I \in \mathcal{E}_{k+l} \setminus \mathcal{E}_{k+l+1}} |a_I| \\
&= \sum_{l=0}^{\infty} \frac{\varepsilon}{(1 + \varepsilon)^{l+1}} \sum_{k \in \mathbb{Z}} (1 + \varepsilon)^{k+l+1} \sum_{I \in \mathcal{E}_{k+l} \setminus \mathcal{E}_{k+l+1}} |a_I| \\
&\geq \sum_{l=0}^{\infty} \frac{\varepsilon}{(1 + \varepsilon)^{l+1}} \sum_{k \in \mathbb{Z}} \sum_{I \in \mathcal{E}_{k+l} \setminus \mathcal{E}_{k+l+1}} |\langle f \rangle_I| |a_I| \quad (\text{by definition of } \mathcal{E}_k) \\
&\geq \sum_{l=0}^{\infty} \frac{\varepsilon}{(1 + \varepsilon)^{l+1}} \sum_{I \in \mathcal{L}} |\langle f \rangle_I| a_I = \left| \sum_{I \in \mathcal{L}} \langle f \rangle_I a_I \right|.
\end{aligned}$$

*Claim 2:*  $\int Mf \geq \varepsilon \sum_{k \in \mathbb{Z}} (1 + \varepsilon)^k |E_{k+1}|$ .

This is done by a commonly used trick

$$\begin{aligned}
\int Mf &= \int_0^{\infty} |\{x : Mf > t\}| dt = \sum_{k \in \mathbb{Z}} \int_{(1+\varepsilon)^k}^{(1+\varepsilon)^{k+1}} |\{x : Mf > t\}| dt \\
&\geq \varepsilon \sum_{k \in \mathbb{Z}} (1 + \varepsilon)^k |\{x : Mf > (1 + \varepsilon)^{k+1}\}| = \varepsilon \sum_{k \in \mathbb{Z}} (1 + \varepsilon)^k |E_{k+1}|.
\end{aligned}$$

To finish, we combine the two claims and conclude that for any  $\varepsilon > 0$ ,

$$\left| \sum_{I \in \mathcal{L}} \langle f \rangle_I a_I \right| \leq (1 + \varepsilon) \|f\|_{H^1_*} \text{Carl}(|a_I|).$$

Finally, letting  $\varepsilon \rightarrow 0$ , we prove the desired inequality. □

#### ACKNOWLEDGEMENT

The author would like to thank his PhD thesis advisor, Serguei Treil, for suggesting this problem for his topics exam and for the invaluable guidance and support throughout the course of preparation.

#### REFERENCES

- [1] B. J. Davis, *On the integrability of the martingale square function*, Israel J. Math. 8, 187-190 (1970)
- [2] A. M. Garcia, *Martingale Inequalities: Seminar Notes on Recent Progress*, New York: Benjamin Inc. 1973, Mathematics Lecture Notes Series
- [3] A. M. Garcia, *The Burgess Davis inequalities via Fefferman's inequality*, Arkiv För Matematik, Volume 11, Numbers 1-2 (1973), 229-237
- [4] C. Fefferman, *Characterizations of Bounded Mean Oscillation*, Bull. Amer. Math. Soc., Volume 77, no. 4 (1971), 587-588
- [5] J. B. Garnett, *Bounded Analytic Functions*, Graduate Texts in Mathematics 236. Springer 1981
- [6] S. Pott and A. Volberg, *Carleson measure and Balayage*, Int. Math. Res. Not. IMRN 2010, no. 13, 2427-2436
- [7] S. Treil, *Commutators, Paraproducts and BMO in non-homogeneous martingale settings*, ArXiv:1007.1210 [math.CA] (2010)

DEPARTMENT OF MATHEMATICS, BROWN UNIVERSITY, PROVIDENCE, RI 02912, USA

E-mail address: jglai@math.brown.edu